

基于周期性多态洗衣机运行的混合事件检测算法

黄 伟, 郑明杨

华北电力大学电气与电子工程学院, 北京
Email: huangwei@ncepu.edu.cn, zmy_18513306974@163.com

收稿日期: 2021年5月15日; 录用日期: 2021年5月29日; 发布日期: 2021年6月10日

摘 要

事件检测是非侵入式负荷监测与识别的重要环节。针对洗衣机运行时存在多事件无法识别的问题, 本文提出了基于周期性多态洗衣机运行的混合事件检测算法, 先采用CUSUM算法捕捉启停波形, 引入shapeDTW算法对其中的周期性波形进行匹配识别, 并将其从总负荷数据中剥离, 避免周期性波形对其他电器事件检测的影响, 该方法能具有很高的识别未知周期性波形的自适应能力。在多态洗衣机参与运行的情况下, 投切各种特性负荷进行仿真分析、方法对比, 该算法保证了特征提取的准确性, 具有很高的检测精度, 为事件检测算法的优化提供了借鉴意义。

关键词

非侵入式负荷监测, 事件检测, 累计和(CUSUM), shapeDTW, 洗衣机

Hybrid Event Detection Algorithm Based on Periodic Polymorphic Washing Machine Operation

Wei Huang, Mingyang Zheng

School of Electrical and Electronic Engineering, North China Electric Power University, Beijing
Email: huangwei@ncepu.edu.cn, zmy_18513306974@163.com

Received: May 15th, 2021; accepted: May 29th, 2021; published: Jun. 10th, 2021

Abstract

Event detection is an important part of non-intrusive load monitoring and identification. Aiming at the problem that multiple events cannot be recognized during the operation of washing machine,

this paper proposes a mixed event detection algorithm based on the operation of periodic polymorphic washing machine. Firstly, the start and stop waveform is captured by the CUSUM algorithm, and the shapeDTW algorithm is introduced to match and identify the periodic waveform, and it is stripped from the total load data. The method can avoid the influence of periodic waveform on the detection of other electrical events and has high adaptive ability to identify unknown periodic waveform. In the case of multi-state washing machine participating in the operation, the simulation analysis and method comparison of various characteristic loads are carried out. The algorithm ensures the accuracy of feature extraction and has high detection accuracy, which provides a reference for the optimization of event detection algorithm.

Keywords

Non-Intrusive Load Monitoring, Event Detection, Cumulative Sum (CUSUM), shapeDTW, Washing Machine

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

负荷监测分为非侵入式和侵入式两种。非侵入式负荷监测[1] (Non-intrusive Load Monitoring, NILM) 旨在监视设备打开、关闭操作、工作时间以及其他相关信息。NILM 将总功耗分解为单个设备的功耗, 与侵入式负荷监测相比具有低成本, 易于安装、拆卸和维护的优点。NILM 主要包含 3 个步骤, 分别是事件检测、特征提取和负荷识别。其中特征提取和负荷识别两部分均以事件检测为基础, 因此, 保证事件检测的准确性是 NILM 监测的关键之处。

目前事件检测主要分为两类: 启发式方法和概率模型方法。第一类算法基于直观或经验, 制定规则对信号突变时刻进行检测。在文献[2]中, 通过计算两个连续样本点之间差值的绝对值来检测总有功功率信号中的事件, 然后将其与预定义阈值进行比较。由于事件检测的多样性和差异性, 依靠直观经验得到的检测结果漏检率、误检率较高, 因而启发式方法应用较少[3]。基于概率模型的事件检测方法检测一个序列中元素的分布或分布参数中某个变化发生的时刻, 本质上为一种统计学判断假设的问题。在 NILM 的背景下, 主要有 4 种方法用于检测电气量信号的变化: 广义似然比(Generalized Likelihood Ratio Test, GLRT)、卡方拟合优度、贝叶斯信息准则(Bayesian Information Criterion, BIC)和累计和(Cumulative Sum, CUSUM)方法。GLRT 是通过从平均值潜在在变化前后的对数概率分布比率中导出一个决策函数来检测新事件的发生[4] [5] [6]。卡方拟合优度是通过判断两个相邻窗口是否共享同一个分布来检测事件[7] [8] [9]。BIC 是通过将观测到的随机样本序列划分成均匀片段, 使用模型复杂度惩罚的似然准则函数进行事件判断[10] [11]。CUSUM 算法是通过判断累计和是否超过阈值来检测事件[12] [13] [14]。两类方法的结合使用可能会产生更好的结果, 文献[15]运用混合算法依次消除了错误事件, 解决了事件检测背景下高波动、瞬态时间长的问题。文献[16]依次使用不同类型的事件检测器来解决“检测噪声”的干扰。

上述事件检测是在“每次最多只有一个负荷发生状态变化”的假设基础上进行的[17], 该假设被称为开关连续性原则 SCP (Switch Continuity Principle) [1]。SCP 与用户内部的负荷使用者数量、负荷数量、采样频率等因素有关[18], 含周期性波形的多个负荷同时发生状态变化时, 负荷特征难以提取, 导致 SCP 不成立, 严重影响负荷分解的结果。如洗衣机、电磁炉和微波炉等。因此, 本文提出了一种针对洗衣机

运行的混合事件检测方法。先采用 CUSUM 算法捕捉启停波形, 引入 shapeDTW 算法对其中的周期性波形进行匹配识别, 并将其从总负荷数据中剥离, 再采用 CUSUM 算法对剥离后的总负荷数据重新检测。建立了自适应周期性波形数据库, 提高了算法对未知周期性负荷的识别。通过实验验证该方法的有效性。

2. 多负荷同时发生状态变化

负荷运行时可能会有多种工作状态, 比如洗衣机洗涤、甩干会产生周期性波形从而产生多个事件, 如图 1 所示。洗衣机一个完整洗涤波形约为 1 s, 间隔约为 1 s。此时打开其他电器, 在 2 s 时间间隔内, 至少有 2 个电器同时发生状态改变。

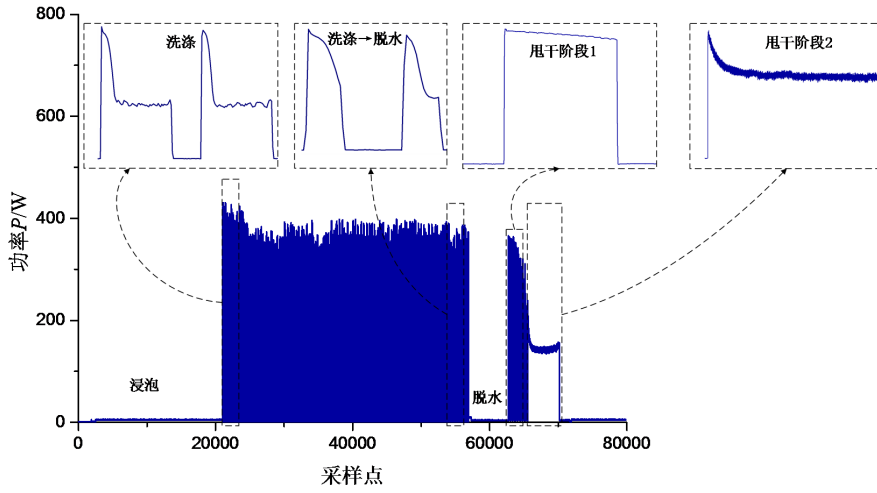


Figure 1. The power curve of washing machine operation
图 1. 洗衣机运行的功率曲线

根据试验分析可知, 电器波形的产生与实际工作机制有关。由于洗衣机是间歇式运行的, 所以会产生周期性波形。基于这一特点, 可通过单独识别周期性波形并剥离, 再重新进行事件检测, 最终避免洗衣机对其他电器的影响。

3. 混合事件检测算法

3.1. CUSUM 事件检测算法

设有一电流有效值(简称电流)时间序列 $I = \{i(j)\}_{j=1}^{\infty}$, 定义两个连续的滑动窗口 W_m (均值计算窗)和 W_d (暂态检测窗), 其长度分别为 m 和 n 。分别计算两个窗口内的均值 M_m 和 M_d , 如式(1)、式(2):

$$M_m = \frac{\sum_{j=k}^{k+m-1} i(j)}{m} \quad (1)$$

$$M_d = \frac{\sum_{j=k+m}^{k+m+n-1} i(j)}{n} \quad (2)$$

式中: k 为时间序列的第一个采样点。

再分别定义 g_k^+ 和 g_k^- 用于检测事件投入和切除, 其计算方式如式(3)、式(4)所示:

$$g_k^+ = \max\{0, g_{k-1}^+ + M_d - (M_m + \beta)\}, g_0^+ = 0 \quad (3)$$

$$g_k^- = \max\{0, g_{k-1}^- - M_d + (M_m - \beta)\}, g_0^- = 0 \quad (4)$$

式中: β 为外界引入的噪声, 表示该电流时间序列电流稳定时的波动水平。

以投入事件为例, 令 $g_0^+ = 0$, 当检测窗口均值 M_d 大于 $M_m + \beta$ 时, g_k^+ 开始大于零。根据检测需求设置累计和阈值 H , 当 $0 < g_k^+ < H$ 时, 说明有事件发生的可能, 但是累计的变化量尚未达到确定事件发生的程度, 此时引入时间延迟因子 d (初始值为 0), 当 $0 < g_k^+ < H$ 时, 令 $d = d + 1$, 计算 g_{k+d}^+ , 直至 $g_{k+d}^+ > H$, 则发生的时刻可根据 $t = k - d$ 倒推得出。为了避免序列震荡导致的事件多识别, 引入时间延迟因子后, 延迟因子每增加 1, 都要对 g_{k+d}^+ 和 g_{k+d-1}^+ 进行比较, 如果 $g_{k+d}^+ \leq g_{k+d-1}^+$, 则认为此时是波动引起的电流变化, 并令 $g_{k+d}^+ = 0, d = 0$, 从而避免了由数据波动导致的多识别事件。以负荷投入事件检测为例的滑动窗双边 CUSUM 事件检测流程见附图 1, 检测切除事件的流程与此相似, 不再累赘。

滑动窗口大小会影响 CUSUM 检测的灵敏度, 窗口大可以有效避开尖峰的影响, 但对样本点不够灵敏, 是因为窗口累计和减小, 导致漏测。应根据不同负荷特征, 设置窗口大小。

3.2. shapeDTW 算法

动态时间规整算法(Dynamic time warping, DTW)是一种在两个时间序列之间搜索最佳比对的方法。它用一个距离来度量测量和样本序列之间的相似性。由于序列在时间维度上出现局部非线性失真, 而采用 DTW 可以很好地解决局部扭曲问题。

设时间序列 \mathbf{X} 、 \mathbf{Y} 的长度分别为 M 、 N , 即 $\mathbf{X} = \{x_1, x_2, \dots, x_M\}$, $\mathbf{Y} = \{y_1, y_2, \dots, y_N\}$ 。构造成对距离矩阵代价矩阵 $\mathbf{C}^{M \times N}$, 其中 $C(i, j)$ 是 x_i 和 y_j 的距离 $d_{ij} = \|y_j - x_i\|_2$, $\|\cdot\|_2$ 表示 2 范数。

设 \mathbf{X} 和 \mathbf{Y} 之间的时间弯曲路径为

$$\mathbf{p} = (p_1, p_2, \dots, p_k, \dots, p_l) \quad (5)$$

其中, $p_k = (i_k, j_k) \in \{1, 2, \dots, i\} \times \{1, 2, \dots, j\}$ 。

计算距离时需要满足边界性, 连续性和单调性三个原则, 如式(6)所示。

$$\begin{cases} p_1 = (1, 1), p_l = (i, j) \\ i_1 \leq i_2 \leq \dots \leq i_l, j_1 \leq j_2 \leq \dots \leq j_l \\ p_{k+1} - p_k \in \{(1, 1), (0, 1), (1, 0)\} \end{cases} \quad (6)$$

两个序列若要得到最佳的匹配, 需要从代价矩阵的起点(1, 1)和终点(M , N)之间找到一条路径来获得最小的累计距离。

根据式(7)所示的递推规则, 用动态规划算法搜索满足上述约束条件的 \mathbf{X} 和 \mathbf{Y} 之间的最优匹配路径, 并计算出相应的最小累计距离 $D(M, N)$, 则 \mathbf{X} 和 \mathbf{Y} 之间的 DTW 距离可以定义为 $DTW(\mathbf{X}, \mathbf{Y}) = D(M, N)/L$, 其中 L 是最优匹配路径的长度。

$$D(i, j) = C(i, j) + \min\{D(i-1, j), D(i-1, j-1), D(i, j-1)\} \quad (7)$$

根据上述原理, DTW 算法在寻找最佳匹配路径的同时对齐两个时间序列的数据点, 使得它们之间的“距离”最小化, 例如图 2 所示。

DTW 算法在某些约束条件下找到了全局最优的匹配, 但它不一定能实现局部合理的匹配。因此在实际应用时会存在一定的误判情况, 原因之一是因为 DTW 算法为寻求全局最优解而过度拉伸和缩放[19], 如附图 2(a)所示。为了解决这个问题, 本文引入了 shapeDTW 算法[20]。shapeDTW 算法先针对每个时序数据点, 利用其前序和后序的部分数据提取出其局部形态特征, 再将每个点周围的局部形状信息结合到动态规划匹配过程中, 从而产生更正确的对齐结果, 即具有相似局部形状的点倾向于被匹配, 而具有不

相似领域的点不太可能被匹配。相比于标准的 DTW 算法只关注全局最优解, shapeDTW 算法更加关注时间序列在局部形态特征上的匹配长度。由此可见, shapeDTW 算法可以利用局部形态特征, 在很大程度上避免了过度拉伸和缩放所导致的误判问题。

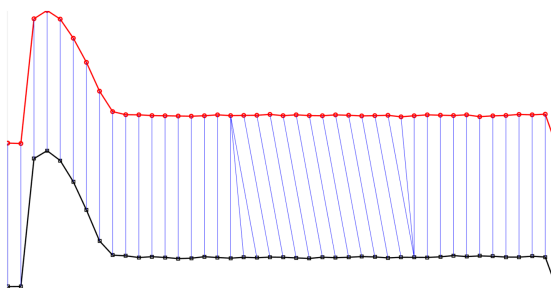


Figure 2. DTW algorithm principle schematic diagram
图 2. DTW 算法原理示意图

3.3. 建立自适应周期性波形数据库

3.3.1. 波形聚类

通过 k-mean 聚类方法对周期性波形进行聚类。考虑到聚类的对象是时间序列, 时间序列可能长度不同。因此采用 DTW 距离代替欧式距离作为样本间的相似度量标准。两个向量 DTW 距离越近则两个向量相似性越高, 划分为同一类的可能性越高, 否则相似性越低, 划分为同一类的可能性越低。假设待聚类的样本集为 $Z = \{z_i | i = 1, 2, \dots, n\}$, 待聚类样本集实际类数为 K 。聚类误差平方和如式(8)所示。

$$E = \sum_{i=1}^K \sum_{j \in C_i} \text{DTW}(z_j - m_i) \quad (8)$$

其中, C_i 是第 i 类的所有样本。 m_i 是第 i 类的聚类中心。 $\sum_{j \in C_i} \text{DTW}(z_j - m_i)$ 表示第 i 类的所有样本到该聚类中心的 DTW 距离之和, 该值越小, 说明该类样本的相似性越高, 聚类效果越好。

为了保证 shapeDTW 算法的有效性, 需建立周期性波形数据库。本文采用改进 k_means 聚类算法对电器的工作波形进行聚类, 将离聚类中心最近的波形作为该电器的典型波形。如图 1 所示, 以洗衣机为例, 洗衣机具有四类典型波形, 设置聚类数量为 4。本文建立的周期性波形数据库共有 6 种波形, 见附图 3。

3.3.2. 自适应周期性波形数据库

为保证周期性波形数据库的构建方法对于多数独立用户具有普适性, 在聚类方法中, 依据周期性负荷具有的操作特性对分类判别条件进行设定, 从而实现周期波形的无监督分类问题。

根据 shapeDTW 的原理只能将待测波形分配到已知的类别中, 但是待测波形可能属于新出现的周期性波形, 该待测波形与已知的类别相似度很小, 因此可以预设一个阈值 T , 来识别新出现的周期性波形。自适应建立波形数据库步骤为: 1) 待测波形 z 与周期性波形数据库的距离满足 $\text{DTW}(z, z_i) \leq T$ 的最小值所对应的类别就判定该波形为哪一类。2) 若 $\text{DTW}(z, z_i) > T$, 则将待测样本暂时判定为新出现的波形类别 z_{new} , 并暂时加入周期性波形数据库。3) 继续匹配下一个波形, 如连续出现 3 个波形都匹配为 z_{new} , 则认为切入了周期性设备 z_{new} , 并永久将 z_{new} 加入周期性波形数据库。反之, 则删除 z_{new} 。

3.4. 混合事件检测算法流程

本文共用到两次 CUSUM 检测, 记为 CUSUM_1 和 CUSUM_2。

1) 输入总负荷电流数据。通常事件检测考虑电流与功率特征, 由于电流数据更加稳定, 本文取电流为被检测的电气量。

2) 检测启停波形。CUSUM₁ 用于检测周期性波形的启停, 由于洗衣进周期性波形的启停间隔较短, 因此需设置较小的窗口。

3) 识别周期性波形。将 2) 中检测到的启停波形进行 shapeDTW 匹配, 若出现连续 3 次波形匹配为同一个电器, 则认为此时打开了该电器。当判断有周期性波形设备打开时, 则将同类波形删除, 即将连续检测到的启 - 停位置之间的数据剥离。最后, 将各个片段重新连接成一个新的时间序列。

4) 检测其他电器波形。CUSUM₂ 用于检测 3) 处理后的数据。此时可认为无周期性电器运行, 即 SCP 成立。因此窗口设置较大。具体设置可参考案例中电器的事件间隔大小。

本文提出的混合事件检测算法流程图如图 3 所示。

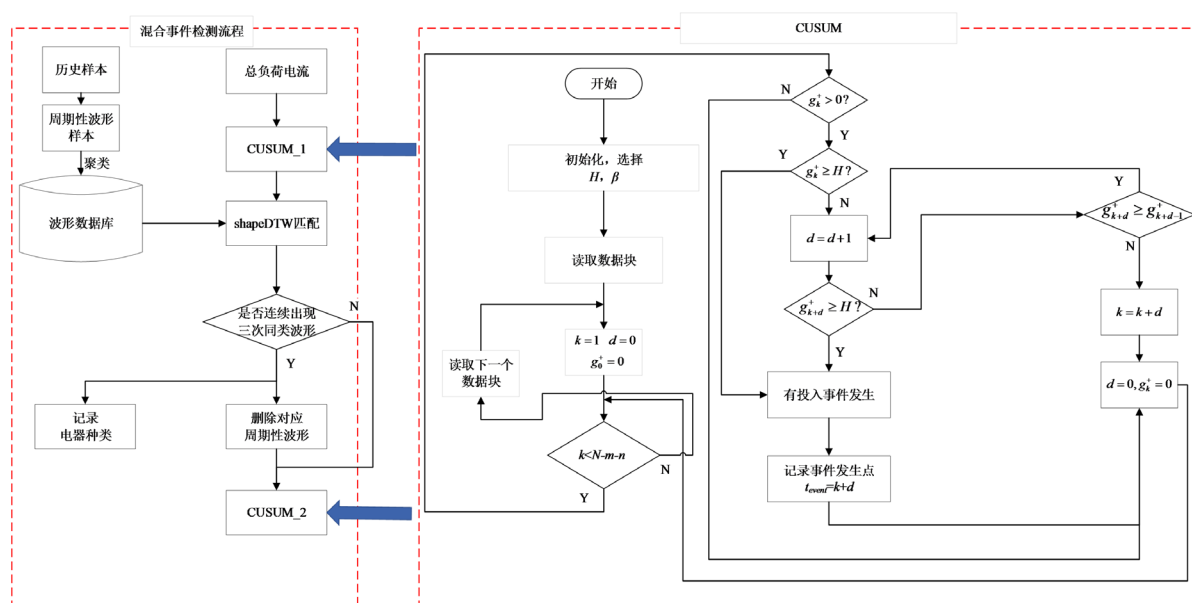


Figure 3. Mixed event detection algorithm flow chart

图 3. 混合事件检测算法流程图

4. 试验与分析

4.1. 参数设置

4.1.1. 窗口设置

CUSUM₁ 负责检测周期负荷事件, 窗口长度设置较小, 取均值计算窗口长度 m 为 2, 检测窗口长度 n 为 2。本算例除洗衣机外, 事件发生间隔大多数在 50 个采样点以上, 因此 CUSUM₂ 窗口长度设置为 50, 取均值计算窗口长度 m 为 40, 检测窗口长度 n 为 10。shapeDTW 算法中形状特征采用一维的梯度直方图[20]。对比算法 CUSUM 和 GLR 算法的窗口均设置为 50。

4.1.1. 阈值设置

设 CUSUM 的阈值为 H , GLR 的阈值为 h 。本节 CUSUM₁、CUSUM₂ 的阈值 H_1 、 H_2 分别设置为 1 和 3.5, 对比算法 CUSUM, GLR 的阈值分别为 $H_3 = 3.5$, $h = 2$ 。shapeDTW 的阈值 T 设置为 185。 H 、 h 、 T 的取值可通过训练获得, 训练过程可参考附录 3。

本节通过 2 个实例的研究来验证所述算法的有效性。实例 1 是在洗衣机运行的基础上, 投入吹风机、电磁炉、电视机、热水壶。测量其事件检测效果和特征误差, 并与当前常用的 CUSUM、GLR 算法进行对比。实例 2 是在洗衣机运行的基础上, 投入 13 种不同特性的电器, 重复算例 1 的试验。

4.2. 评价指标

事件检测可以看成是一个二分类问题, 即对每一时刻的检测结果只有两种可能: 发生或未发生负荷事件。因此引入二分类算法性能评价中常用的混淆矩阵(Confusion Matrix), 将检测结果分成四类: 真正类(TP, 实际的负荷事件被检测出)、假正类(FP, 未发生负荷事件却被误检为有发生)、假负类(FN, 实际发生的负荷事件被漏检)和真负类(TN, 实际未发生负荷事件, 且与检测结果一致)。在进行有限次负荷投切操作的情况下, TN 的值一定是远大于其他三类的。为了规避这种样本类别失衡的影响, 选取的评估指标应不依赖于 TN 类的值。因此, 以精确率(Precision)、召回率(Recall)和 F_Score 作为测试算例结果的评价指标, 三种评价指标可分别按照式(9)、式(10)和式(11)计算求得。

$$Precision = \frac{|TP|}{|TP| + |FP|} \tag{9}$$

$$Recall = \frac{|TP|}{|TP| + |FN|} \tag{10}$$

$$F_Score = \frac{2}{\frac{1}{Precision} + \frac{1}{Recall}} \tag{11}$$

如式(9)和式(10)所示, 精确率表示所有被检测为发生的负荷事件中实际发生的事件的占比, 其衡量事件检测方法能否检测准确的能力; 召回率表示检测出的负荷事件占所有实际发生的事件的比例, 其衡量事件检测方法能否检测全面的能力。如式(11)所示, F_Score 是精确率和召回率的调和平均值, 其能综合衡量事件检测算法在精确率和召回率这两个指标上的得分。

4.3. 实例 1

本算例采用用户在某个时间段的电流数据作为算例, 如图 4(a)所示, 共包含 167,585 个采样点, 采样

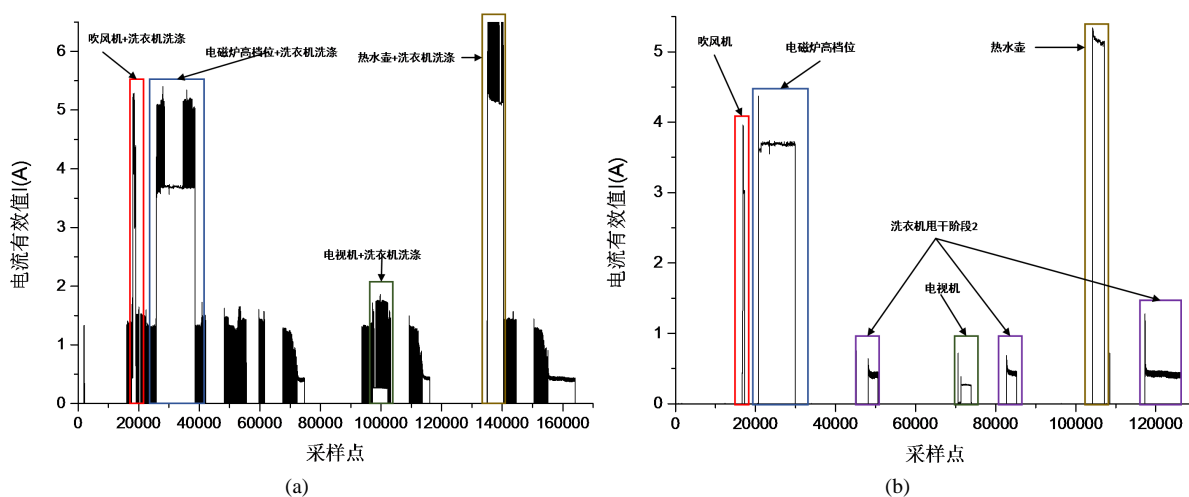


Figure 4. RMS curve of mixed operation current of washing machine, hair dryer, induction cooker, TV set and hot kettle
图 4. 洗衣机、吹风机、电磁炉、电视机、热水壶混合运行电流有效值曲线

频率为 50 Hz, 包括洗衣机、吹风机、电磁炉、电视机、热水壶 5 个电器, 除洗衣机外共有 10 个负荷事件。为测试事件检测对特征提取的影响, 分别对每次事件提取功率特征。表 2 为吹风机、电磁炉、电视机、热水壶单独运行时的特征。

如图 4(b)所示, 经过 3.4 节的第 3 步处理后, 洗衣机运行所产生的周期性波形几乎都被删除, 仅剩下洗衣机甩干 2 阶段的波形与其他电器运行数据。第 3 步处理过程见附录 2。再进行 3.4 第 4 步处理, 结果如表 1 所示。根据表 1 提取增量特征, 结果如表 3 所示。

Table 1. Results of event detection

表 1. 事件检测结果

电器	事件检测属性	本文算法	CUSUM	GLR
吹风机	打开(低档)	√	/	√
	低档→高档	√	√	√
	高档→中档	√	√	√
	关闭	√	√	√
电磁炉	打开	√	√	√
	关闭	√	√	√
电视机	打开	√	√	/
	关闭	√	√	√
热水壶	打开	√	√	√
	关闭	√	√	√

Table 2. Operating power of hair dryer, induction cooker, TV set and hot kettle

表 2. 吹风机、电磁炉、电视机、热水壶运行功率

电器	事件检测属性	特征提取/功率
吹风机	低档	82.4189919
	低档转中档	375.9587837
	中档转高档	376.337626
	低档转高档	756.789123
	低档关闭	-80.113430
	中档关闭	-455.990558
	高档关闭	-833.334390
电磁炉高档位	打开	809.7102178
	关闭	-806.390558
电视机	打开	53.0343093
	关闭	-50.975868
热水壶	打开	1173.465605
	关闭	-1121.69973

Table 3. Comparison of feature extraction results
表 3. 特征提取结果对比

	电器 事件检测 属性	吹风机			电磁炉		电视机		热水壶		
		打开 (低档)	低档→ 高档	高档→ 中档	中档 关闭	打开	关闭	打开	关闭	打开	关闭
功率 特征 (W)	本文算法	81.07	774.33	-378.97	-456.42	809.09	-811.74	52.14	-53.84	1176.58	-1115.76
	CUSUM	-	764.79	-658.21	-477.92	1154.65	-1106.88	415.15	-49.13	1454.42	-1504.71
	GLR	456.36	764.77	-651.13	-478.15	1155.11	-1106.87	-	-50.15	1454.42	-1504.71
特征 相似度 (%)	本文算法	0.0164	0.0232	0.007	0.0009	0.0008	0.0066	0.0169	0.0562	0.0027	0.0053
	CUSUM	-	0.0106	0.749	0.0481	0.426	0.3726	6.828	0.0362	0.2394	0.3415
	GLR	4.5371	0.0105	0.7302	0.0486	0.4266	0.3726	-	0.0162	0.2394	0.3415

由表 1、表 2 可知, 本文算法与 CUSUM、GLR 在事件的检测效果相似, 但在电器特征提取的准确性上表现更优。由表 3 可知, 基于本文算法提取的特征与电器单独运行时特征相近, 而基于 CUSUM 与 GLR 算法所提取的特征与电器单独运行时特征存在较大的差距。通过分析发现: CUSUM 和 GLR 所检测到的事件包含洗涤波形, 如图 5 所示。CUSUM 算法与 GLR 算法即使检测到该事件, 特征提取会包含 2 个电器的特征叠加, 导致负荷识别错误。而本文算法, 是在剥离周期性波形的基础上进行事件检测, 因此不存在特征叠加问题。因此, 本文算法可保证特征提取以及负荷识别的准确性。由于删除了洗衣机波形, 检测到的时间点与实际事件发生的时间点相差大约 1 s, 即一个波形的时间。

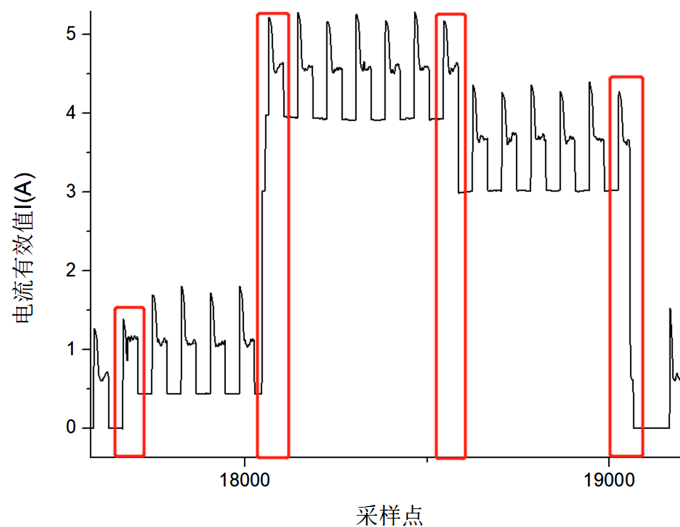


Figure 5. The current effective value curve of hair dryer and washing machine in mixed operation
图 5. 吹风机与洗衣机混合运行的电流有效值曲线

4.4. 实例 2

本算例以仿真实验室连续测量十次洗衣机运行的电流数据作为算例, 如图 6 所示, 共包含 1,800,000 个采样点, 采样频率为 50 Hz, 包括洗衣机、吹风机、电磁炉、电视机、热水壶、电饭煲、电风扇、热水器、消毒柜、油烟机、定频空调、微波炉、电暖气、电烤箱 14 个电器, 除洗衣机外共有 73 个负荷事件。事件检测结果如表 4 所示。

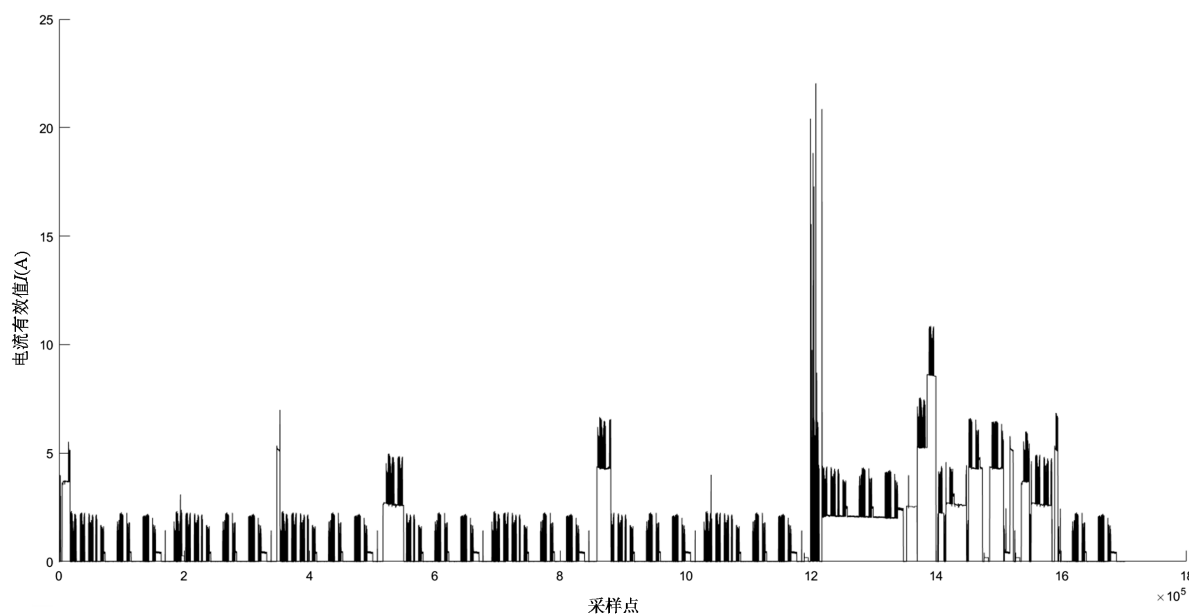


Figure 6. Current effective value curve

图 6. 有效值曲线

Table 4. Results of event detection

表 4. 事件检测结果

	本文算法	CUSUM	GLR
事件数	74	70	69
误检	4	5	7
漏检	3	8	11
TP	70	65	62
FP	4	5	7
FN	3	8	11
精确率	0.9459	0.9286	0.8986
召回率	0.9589	0.8904	0.8493
F_score	0.9524	0.9091	0.8732

由表 4 可得, 本文算法在精确率上与 CUSUM 算法相似, 但在召回率和 F_score 上的表现更优(分别提高了 6.85% 和 4.33%)。而与 GLR 算法相比, 本文算法在三个指标上都有更优的表现(分别提高了 4.74%、10.96% 和 7.91%)。基于上述分析, 在含有洗衣机等周期性电器的家庭中, 本文提出的基于 CUSUM 与 shapeDTW 混合事件检测算法具有比 CUSUM 与 GLR 更好的事件检测效果, 更能保证特征提取的准确性和有效性。

5. 结论

事件检测是非侵入式负荷监测的基础环节, 其准确性将影响整个负荷分解与识别的效果。针对洗衣机运行造成的多事件同时发生的问题, 本文研究了混合事件检测方法。通过仿真实验数据验证对比表明:

1) 在洗衣机运行背景下, 就对特征提取的影响而言, 相比于 CUSUM 和 GLR 算法, 混合事件检测算法可保证提取的增量特征的准确性, 为后续的特征提取、负荷辨识等工作打下了基础。

2) 在洗衣机运行背景下, 就检测效果而言, 混合事件检测算法较 CUSUM 和 GLR 算法具有更高的精确率。因此更适用于用户非侵入式负荷监测的工程实际。

本文算法不仅适用于洗衣机, 还适用于电磁炉、微波炉等其他含有周期性波形的设备的识别。同时还具备识别未知周期性设备和自适应建立周期性波形数据库的能力。综上, 本文算法在提高检测能力的同时, 可以保证特征提取的准确性, 提高负荷识别准确率。

参考文献

- [1] Hart, G.W. (1992) Nonintrusive Appliance Load Monitoring. *Proceedings of the IEEE*, **80**, 1870-1891. <https://doi.org/10.1109/5.192069>
- [2] Weiss, M., Helfenstein, A., Mattern, F. and Staake, T. (2012) Leveraging Smart Meter Data to Recognize Home Appliances. 2012 *IEEE International Conference on Pervasive Computing & Communications*, Lugano, 19-23 March 2012, 190-197. <https://doi.org/10.1109/PerCom.2012.6199866>
- [3] 张露, Auger, F., 荆朝霞, Houidi, S., Bui, H.K., 肖江. 基于非侵入式的事件检测方法统计评估[J]. 电测与仪表, 2020, 57(1): 106-112, 120.
- [4] Berges, M., Goldman, E., Scott Matthews, H. and Soibelman, L. (2009) Learning Systems for Electric Consumption of Buildings. *Proceedings of 2009 ASCE International Workshop on Computing in Civil Engineering*, Austin, 24-27 June 2009, 1-10. [https://doi.org/10.1061/41052\(346\)1](https://doi.org/10.1061/41052(346)1)
- [5] Yang, C.C., Soh, C.S. and Yap, V.V. (2014) Comparative Study of Event Detection Methods for Non-Intrusive Appliance Load Monitoring. *Energy Procedia*, **61**, 1840-1843. <https://doi.org/10.1016/j.egypro.2014.12.225>
- [6] Tsai, M. and Lin, Y. (2012) Modern Development of an Adaptive Non-Intrusive Appliance Load Monitoring System in Electricity Energy Conservation. *Applied Energy*, **96**, 55-73. <https://doi.org/10.1016/j.apenergy.2011.11.027>
- [7] De Baets, L., Ruysinck, J., Develder, C., Dhaene, T. and Deschrijver, D. (2017) On the Bayesian Optimization and Robustness of Event Detection Methods in NILM. *Energy and Buildings*, **145**, 57-66. <https://doi.org/10.1016/j.enbuild.2017.03.061>
- [8] Jin, Y., et al. (2011) A Time-Frequency Approach for Event Detection in Non-Intrusive Load Monitoring, SPIE Defense, Security, and Sensing. *International Society for Optics and Photonics*, 80501U-80501U. <https://doi.org/10.1117/12.884385>
- [9] Yang, C.C., Soh, C.S. and Yap, V.V. (2015) A Systematic Approach to ON-OFF Event Detection and Clustering Analysis of Non-Intrusive Appliance Load Monitoring. *Frontiers in Energy*, **9**, 231-237. <https://doi.org/10.1007/s11708-015-0358-6>
- [10] 肖江, Auger, F., 荆朝霞, Houidi, S. 基于贝叶斯信息准则的非侵入式负荷事件检测算法[J]. 电力系统保护与控制, 2018, 46(22): 8-14.
- [11] 陈中, 方国权, 赵家庆, 丁宏恩. 基于贝叶斯迭代的非侵入式负荷事件检测方法[J]. 电测与仪表, 2021, 58(4): 1-8.
- [12] Zhu, Z., Zhang, S., Wei, Z., Yin, B. and Huang, X. (2018) A Novel CUSUM-Based Approach for Event Detection in Smart Metering. *IOP Conference Series: Materials Science and Engineering*, **322**, Article ID: 072014. <https://doi.org/10.1088/1757-899X/322/7/072014>
- [13] 牛卢璐, 贾宏杰. 一种适用于非侵入式负荷监测的暂态事件检测算法[J]. 电力系统自动化, 2011, 35(9): 30-35.
- [14] 史帅彬, 张恒, 邓世聪, 周东国, 周洪, 胡文山. 基于复合滑动窗的 CUSUM 暂态事件检测算法[J]. 电测与仪表, 2019, 56(17): 13-18.
- [15] Lu, M. and Li, Z. (2019) A Hybrid Event Detection Approach for Non-Intrusive Load Monitoring. *IEEE Transactions on Smart Grid*, **11**, 528-540. <https://doi.org/10.1109/TSG.2019.2924862>
- [16] 徐志翔. 基于机器学习的非侵入式负荷监测技术研究[D]: [硕士学位论文]. 杭州: 浙江大学, 2020.
- [17] 郭红霞, 陆进威, 杨苹, 刘泽健. 非侵入式负荷监测关键技术问题研究综述[J/OL]. 电力自动化设备, 2021, 41(1): 135-146. <https://doi.org/10.16081/j.epae.202011001>, 2020-12-21
- [18] Makonin, S. (2016) Investigating the Switch Continuity Principle Assumed in Non-Intrusive Load Monitoring(NILM). 2016 *IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, Vancouver, 15-18 May 2016,

- 1-4. <https://doi.org/10.1109/CCECE.2016.7726787>
- [19] 杨子元, 许晓斌, 李欣, 赵一萌. 基于智能感知技术的用电事件识别方法研究[J]. 物联网学报, 2019, 3(4): 109-115.
- [20] Zhao, J. and Itti, L. (2017) shapeDTW: Shape Dynamic Time Warping. *Pattern Recognition*, **74**, 171-184. <https://doi.org/10.1016/j.patcog.2017.09.020>

附录

1. 正文图

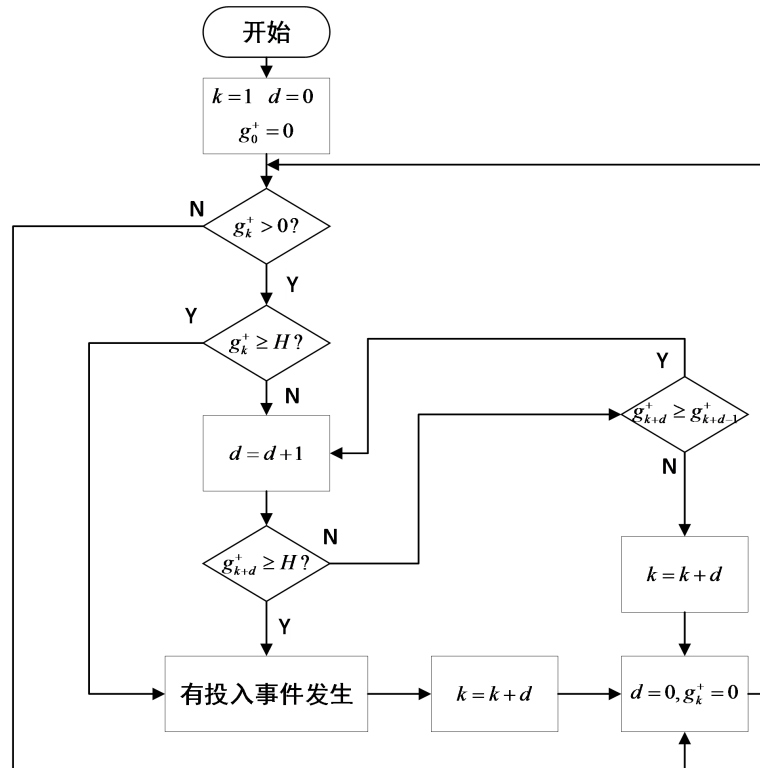


Figure A1. Sliding window bilateral CUSUM event detection algorithm flow
附图 1. 滑动窗双边 CUSUM 事件检测算法流程

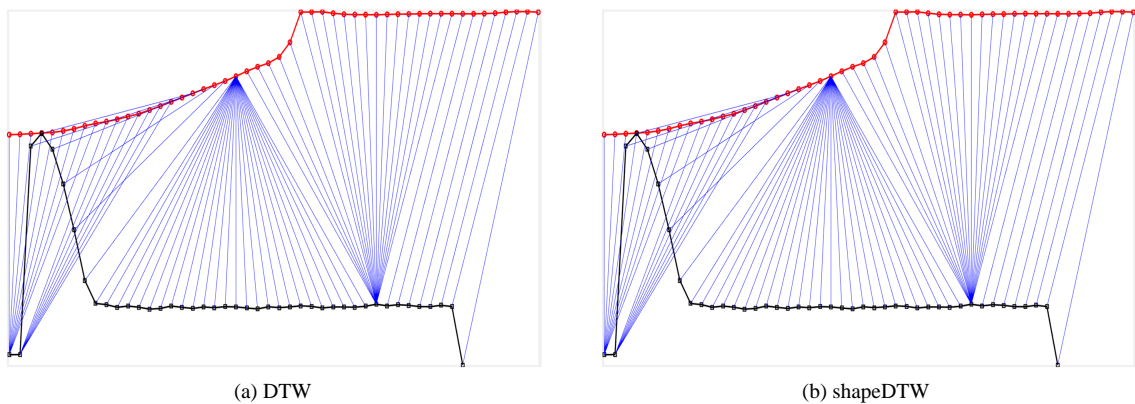


Figure A2. Comparison of DTW and SHAPEDTW matching paths
附图 2. DTW 与 shapeDTW 匹配路径对比

2. 建立自适应波形数据库

1) 周期性波形数据库

波形数据库中一共 6 种波形, 如附图 3 所示: ① 洗衣机洗涤; ② 洗衣机洗涤→脱水; ③ 洗衣机甩干 1 阶段; ④ 电磁炉文火挡位; ⑤ 电饭煲保温挡位; ⑥ 微波炉低档位。

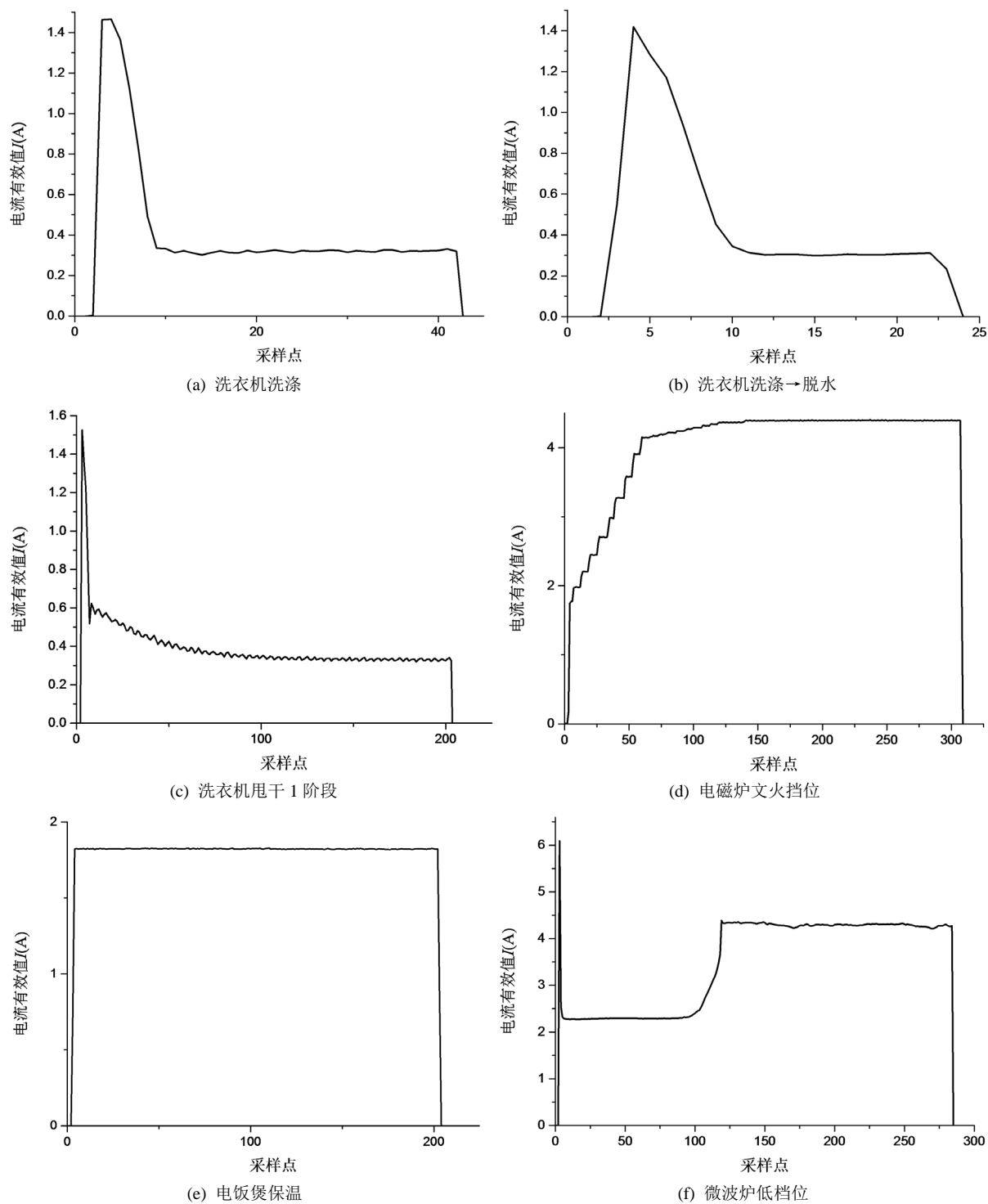


Figure A3. Periodic waveform database

附图 3. 周期性波形数据库

2) 检测周期性波形流程

- ① 通过 CUSUM_1 检测完整的启停波形, 如附图 4 所示。
- ② 将检测到的波形进行 shapeDTW 匹配, 以附图 4 波形为例, shapeDTW 匹配结果如附表 1 所示。

③ 若出现连续 3 次波形匹配为同一个电器, 则认为此时打开了该电器。

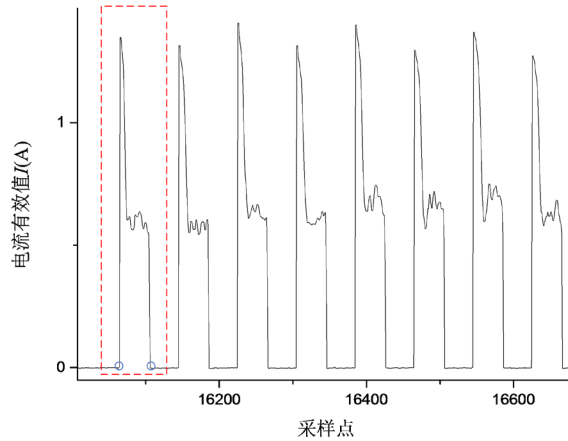


Figure A4. Start and stop waveform detection in washing machine
附图 4. 洗衣机洗涤阶段启停波形检测

Table A1. Matching results of shapeDTW

附表 1. shapeDTW 匹配结果

波形	洗衣机洗涤	洗衣机洗涤→脱水	洗衣机甩干 1 阶段	电磁炉文火挡位	电饭煲保温	微波炉低档位
DTW 距离	1.58	19.5	74.96	50.78	42.58	290.73

由附表 1 可知, 待匹配波形与洗衣机洗涤的 DTW 距离最小, 即该波形属于洗衣机洗涤波形。

3. 阈值训练

1) H 、 h 训练过程

考虑到本算例中, 除洗衣机外, 电器的事件发生间隔大多数在 50 个采样点以上, 取 CUSUM 和 GLR 算法窗口长度为 50。

如附图 5 所示, 共包含 150,000 个采样点, 采样频率为 50 Hz, 包括吹风机、电磁炉、电视机、热水壶、油烟机、消毒柜、热水器、电风扇、电饭煲 9 种电器。检测过程累积和如附图 6 所示。取 $H \in (0, 4)$, 间隔为 0.5。取 $h \in (0, 4)$, 间隔为 1。事件检测结果如附表 2、附表 3 所示。

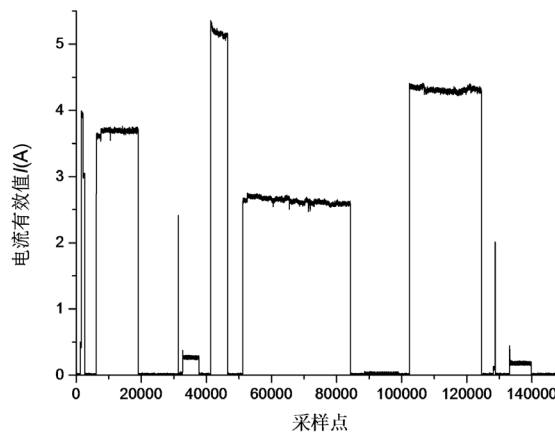


Figure A5. Current effective value curve
附图 5. 电流有效值曲线

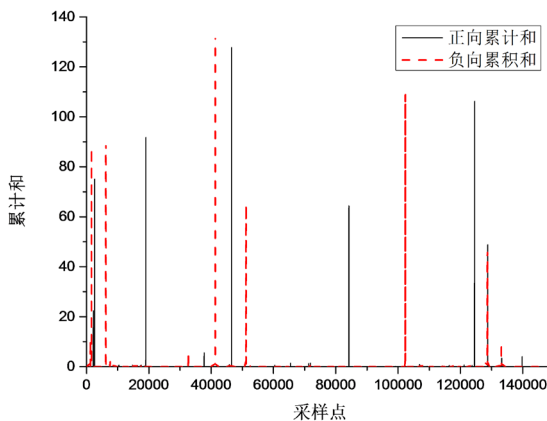


Figure A6. Test process accumulation and schematic diagram
附图 6. 检测过程累积和示意图

Table A2. Event detection results of CUSUM
附表 2. CUSUM 事件检测结果

H	事件数	误检	漏检	TP	FP	FN	精确率	召回率	F_score
0.5	41	22	0	19	22	0	0.46	1.00	0.63
1	25	6	0	19	6	0	0.76	1.00	0.86
1.5	22	3	0	19	3	0	0.86	1.00	0.93
2	21	2	0	19	2	0	0.90	1.00	0.95
2.5	19	1	1	18	1	1	0.95	0.95	0.95
3	19	1	1	18	1	1	0.95	0.95	0.95
3.5	18	0	1	18	0	1	1.00	0.95	0.97
4	18	0	1	18	0	1	1.00	0.95	0.97
4.5	17	0	2	17	0	2	1.00	0.89	0.94

Table A3. Event detection results of GLR
附表 3. GLR 事件检测结果

h	事件数	误检	漏检	TP	FP	FN	精确率	召回率	F_score
1	18	1	3	18	1	3	0.95	0.86	0.90
2	17	0	3	17	0	3	1.00	0.85	0.92
3	16	0	3	16	0	3	1.00	0.84	0.91
4	16	0	3	16	0	3	1.00	0.84	0.91
5	16	0	4	16	0	4	1.00	0.80	0.89

附表 2 可得, H 取 3.5 时, CUSUM 检测效果最好。由附表 3 可得, h 取 3.5 时, GLR 检测效果最好。

2) T 的训练过程

T 的取值应至少满足: 同类波形到聚类中心的 DTW 距离 T_r 小于 T 。如式(12)所示。

$$T = \max(T_r), r = 1, 2, 3, 4, 5, 6$$

$$T_r = \max_{j \in C_i} (\text{DTW}(z_j - m_i)) \tag{12}$$

其中 C_i 是第 i 类的所有样本。 m_i 是第 i 类的聚类中心。 $DTW(z_j - m_i)$ 表示第 i 类的样本 z_j 到该聚类中心的 DTW 距离。

Table A4. T_r distances for different waveforms

附表 4. GLR 事件检测结果

波形	波形 1	波形 2	波形 3	波形 4	波形 5	波形 6
T_r	94.82	7.85	184.25	100.04	4.51	85.00

本文周期性波形数据库共有 6 种波形, T_r 如附表 4 所示。因此 $T \geq 184.25$, 取 $T = 185$ 。